Supplemental Online Material for "Reconciling disparate 20th Century Indo-Pacific ocean temperature trends in the instrumental record" by A. Solomon and M. Newman

In this supplementary material we provide a detailed discussion of how the linear inverse model and optimal perturbation filter were constructed for the paper, as well as, a series of diagnostics[10,11,17,25,26,27] to demonstrate the robustness of the results presented in the paper.

1) Parameters used to construct the LIM

The optimal perturbation filter has 5 free parameters; the truncation of the data in EOF space, the lag used to calculate the lag-covariance statistics (τ_0), the interval over which the growth of the optimal structure is maximized (τ_e), the time over which the filter is applied (τ_1), and the norm under which the propagator **G** is optimized. In addition, the time period used to construct the lag covariance statistics needs to be specified. For the SODA data we use the 50-year period between 1957-2006 to construct the LIM. For the SST datasets we use the 120-year common period of 1891-2010 to construct the LIM from each dataset.

Prefiltering (truncation) in EOF space and the parameter τ_0 are used to construct the LIM. We use $\tau_0 = 3$ months for the lag-covariance statistics, similar to other studies[10,11,17,25,26,27]. For the LIM to be a valid approximation of the system's dynamics it is necessary to demonstrate that it can reproduce observed lag-covariance statistics at much longer lags than τ_0, i.e. the tau-test[17]. The linearity of the system's dynamics, and the validity of the LIM, is demonstrated by the similarity between LIM and observed lag-covariance statistics at 9 and 18 months, (shown for the HadISST data in Supp. Fig. 2).

The impact of the EOF truncation on the forecast skill of the LIM is demonstrated by sub-sampling the data record by sequentially removing one ten-year period, computing **G** for the remaining years, and then generating forecasts for the independent years. This procedure is repeated for the entire period. Forecast skill is then determined by comparing the local anomaly correlation between the cross-validated model

predictions and gridded *untruncated* verifications. The impact of EOF truncation for HadISST SSTs averaged in the NINO3.4 region (4°S-4°N,170°E-120°W) is shown in Supp. Fig. 3. It is seen that increasing the number of EOFs beyond 20 results in a marginal increase in skill for lags greater than 12 months. All LIMs in this paper are constructed with data that has been truncated to 20 EOFs. It is important to note that all trends use the untruncated data where the first 20 EOFs are unfiltered or filtered using the LIM.

2) Parameters used to construct the optimal perturbation filter

The optimal perturbation for SST-LIMs has typically been defined under the L2 norm (i.e., domain-mean squared amplitude), and we adopt this convention as well. The parameters τ_1 and τ_e are chosen to optimize the filter as follows. Based on the maximum possible anomaly growth factor, sometimes called the ‘‘maximum amplification’’ (MA) curve (shown in Supp. Fig. 4a for HadISST), it is seen that the largest possible SST amplification occurs at 9 months, which would then seem to be a reasonable choice for τ_e. However, the evolution from the optimal to peak ENSO is not substantially different for shorter optimization times. For example, the 30-month evolution of EOF1 from the dominant initial conditions (Supp. Fig. 4b) for $\tau_e = 9$ months is similar to that for $\tau_e = 3$ and 6 months, and the optimal initial conditions (Supp. Fig 5) also display somewhat subtle differences. To determine which optimization time is most relevant to actual anomaly growth we look for the optimal that not only grows but is also excited by noise (since an initial condition which *can* grow but is less well forced would have less relevance). We thus determine growth “expectation values” as a function of τ_e, where the expectation values are calculated as the standard deviation of the projection of the optimals for $\tau_e = 3,6,9$ on the noise and then weighted by the singular value. The integrated noise $\boldsymbol{\sigma}(t)$ (i.e., forecast error) is defined from the LIM as

$$\boldsymbol{\sigma}(t) = \mathbf{x}(t) - \mathbf{G}(3)\mathbf{x}(t-3) \quad .$$

The expectation values, shown in Supp. Fig. 6, peak at $\tau_e = 3$, indicating that the most relevant optimal is for $\tau_e = 3$; that is, although the $\boldsymbol{\phi_1}(3)$ optimal doesn’t amplify quite as much as $\boldsymbol{\phi_1}(6)$ and $\boldsymbol{\phi_1}(9)$ it is more forced by the noise.

To choose τ_1, we choose the time at which EOF1 decays to its initial value (from Supp. Fig. 3b). This time is at approximately 24 months for $\tau_e = 9$, but for $\tau_e = 3-6$ it is a little shorter, ~21 months.

Sensitivity to the choices of τ_e and τ_1 is illustrated in Supp. Fig. 7a, which shows the root mean square error (RMSE) of the residual as defined in eq. (4) minus the 3-month forecast error [RMSE($\mathbf{R} - \boldsymbol{\sigma}(t)$)] as a function of EOF for HadISST for the 1900-2010 time period (also shown for ERSST in Supp. Fig. 7b). RMSE($\mathbf{R} - \boldsymbol{\sigma}(t)$) is a measure of how much of the predictable ENSO variability explained by the LIM is removed by the filter. Supp. Fig. 7 shows RMSE for $\tau_1 = 21,24,27$ months with black, blue, and red markers, respectively. The largest impact of the filter is on the amplitude of EOF1. For example, for $\tau_e = 3$ months and $\tau_1 = 21$ months the filtering reduces the RMSE of EOF1 by 73% while reducing the RMSE of EOF2 by less than 6%. The RMSE of EOF1 decreases for decreasing τ_e and increasing τ_1, due to a somewhat larger projection of the optimal initial conditions on EOF1 as τ_e is decreased (Supp. Fig 5).

It is important to note that the dominant initial conditions $\boldsymbol{\phi_1}$ have a little projection on EOF1. This projection is larger for $\tau_e = 3$ months than for $\tau_e = 6$ and 9 months. This is an issue since a number of climate model projections of the response to an increase in greenhouse gases produce a warming pattern in the tropical Indo-Pacific that is "El Niño-like"[32]. This can cause a projection of the trend onto the optimal initial conditions, which may cause a bias in the filtered trends. We test this by comparing filtered trends and time series using $\tau_e = 3,6,9$ months (Supp. Figs. 8-15) and conclude that there is no detectable projection of a systematic trend on the optimals, but this is an issue that needs to be addressed when applying this filter to climate model simulations where there may be a significant projection of the response to external forcing on the optimal initial conditions estimated using SSTs. As discussed in ref. 34, using an EOF basis defined using three dimensional ocean temperatures can minimize this problem since this exploits the fact that ENSO variability has a three dimensional structure that is distinct from the response to external forcing.

Since aliasing due to the large interannual variability on the relatively small long-

term trend makes it impossible to estimate whether this is indeed the case for the datasets used in this paper (i.e., the motivation for this paper), we plot in Supp. Figs. 8-11 the sensitivity of the filtered cold tongue SSTs to τ_e and τ_1 for all four datasets. First, it is interesting to note that the cold tongue SSTs are relatively insensitive to τ_1, i.e., the blue, green, and red lines are indistinguishable. Also note that for all of the datasets increasing τ_e reduces the variability removed from each time series. To see what impact this sensitivity to τ_e has on the long-term trends, we plot in Supp. Figs. 12-15 the 1900-2010 SST trends for all four datasets as a function of τ_e (with τ_1 set to 21 months). The impact is negligible on the HadISST and KAPLAN trends. For the COBE and ERSST datasets, it is seen that the less efficient filtering as τ_e increases, increases the magnitude of the cooling trend in the equatorial Pacific.

Based on the results presented in this supplementary material we use the values of $\tau_e = 3$ months and $\tau_1 = 21$ months to filter the SST datasets in the main body of the paper. LIMs constructed using lag-covariance statistics from different periods produce slightly different dominant evolving structures. It was found that $\tau_1 = 18$ months was optimal for the SODA 1958-2007 LIM rather than the $\tau_1 = 21$ months used with the 1891-2010 SST LIMs.

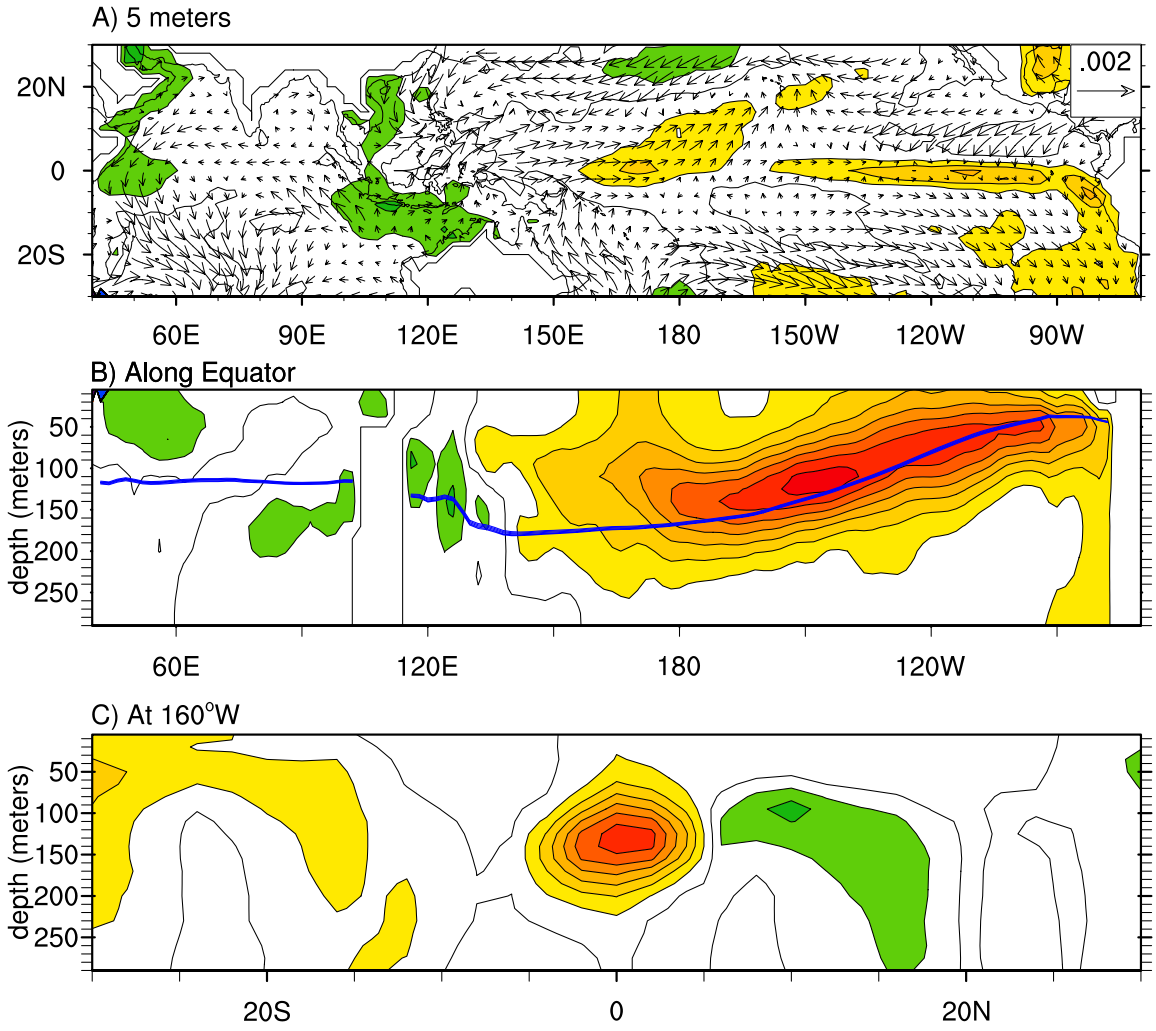
In this study we use the 1891-2010 common period to construct the LIM for each dataset. To justify this we need to demonstrate that the dynamical operator (or deterministic feedback matrix) in eq. (1) does not change appreciably over this period. This question was addressed in an earlier study[11] where it was argued that the dynamical operator has not changed since 1871 but that the earlier part of the record cannot be used to estimate the deterministic feedback matrix because of errors due to sparsity of the data prior to 1948. A finding that motivated this conclusion was that the dominant optimal initial conditions for LIMs constructed with statistics from HadISST over the 1872-1947 and 1949-2004 periods using a $\tau_e = 8$ months had a pattern correlation of 0.25. To investigate this for the datasets and periods used in this study we calculated the pattern correlation of dominant optimal initial conditions for the three periods 1891-1950, 1951-2010, and 1891-2010 using a $\tau_e = 3$ months (correlations shown in Table 1). Similar to the earlier study the pattern correlation between the earlier and later periods is relatively

low for the HadISST data, 0.48 for the 60-year periods. But interestingly, this is not the case for the other three datasets, with correlations ranging from 0.70-0.78. In addition, we find that pattern correlations can be low due to uncertainties in estimates of dominant initial conditions from LIMs that use too little data. To illustrate this point we constructed 1000 synthetic 120-year time series from the 1891-2010 HadISST LIM. Therefore, these time series had the same lag-covariance and noise statistics as the HadISST 1891-2010 record. We divided each time series into two 60-year segments and constructed a LIM with the lag-covariance statistics for each segment. We then calculated the pattern correlation of the dominant optimal initial conditions for each 60-year LIM. It was found that 900 out of the 1000 time series had pattern correlations less than 0.86 and the mean of the correlations was 0.74. This procedure was repeated using 1000 240-year time series divided into two 120-year segments and it was found that 900 out of the 1000 time series had pattern correlations less than 0.93 and the mean of the correlations was 0.86. From this we surmise that, at least for the datasets other than HadISST, it cannot be concluded that the differences in the optimals for the 60-year periods are primarily due to errors in the statistics during the earlier part of the record. Clearly this is an issue, but based on the correlations shown in Table 1 we chose to construct the statistics using the full 120-year period. Further, we found that the results presented in the paper are not changed when the filter uses the evolution of the dominant optimal initial conditions from a LIM constructed with 1951-2010 statistics.

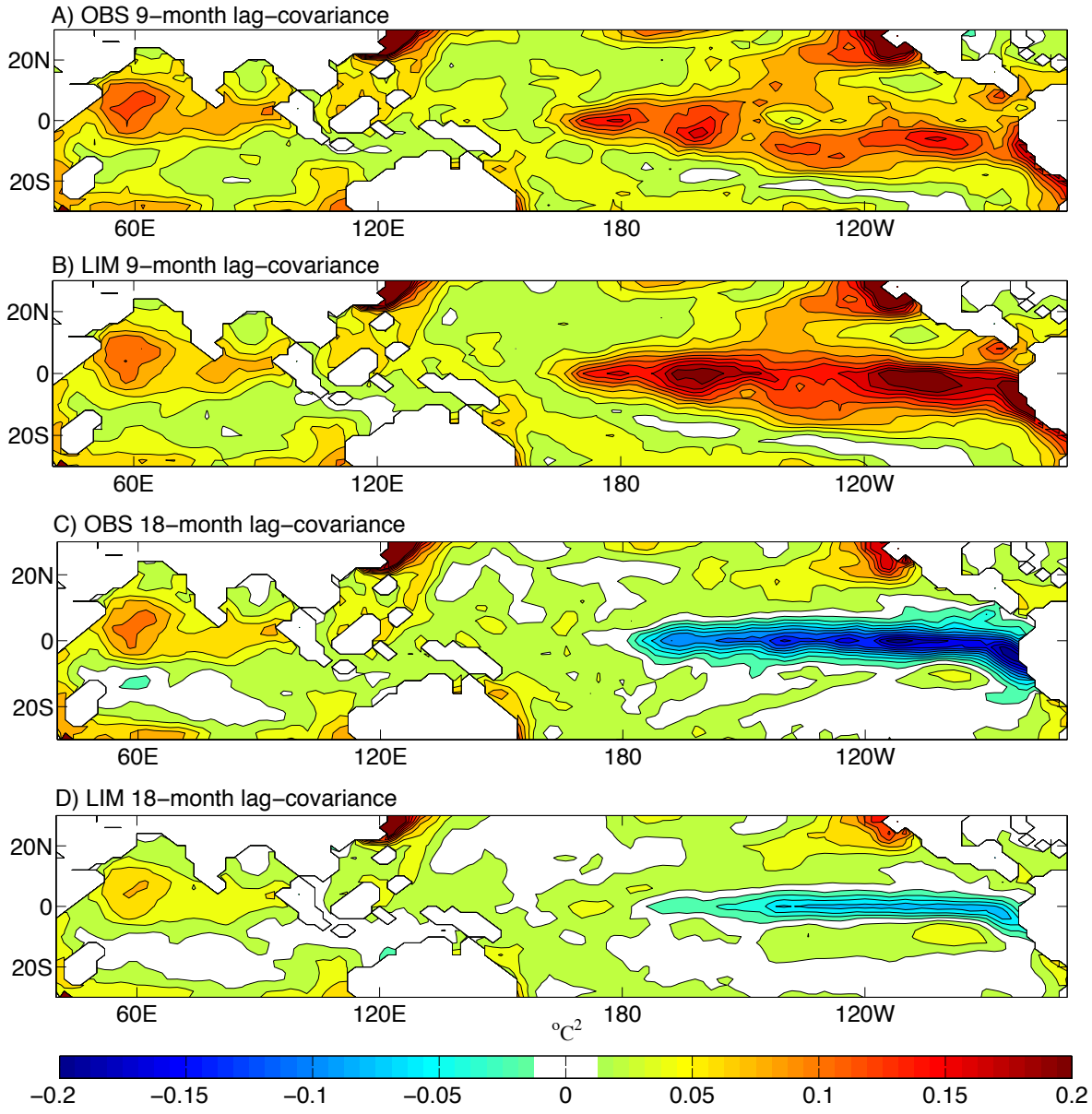
As an additional test of the robustness of the filter we calculated the correlation between the optimal initial conditions for SST amplification over a 3-month interval and the optimal evolved SST structure 6 months later for the 1891-1950 and 1951-2010 period separately. For both periods and for all four datasets the correlations range was 0.68-0.78 indicating that the linear relationship between the initial and evolved structures expressed in the LIM is valid for both 60-year periods.

	HadISST	**ERSST**	**COBE**	**KAPLAN**
cr(1891-1950 , 1891-2010)	0.60	0.82	0.90	0.87
cr(1951-2010 , 1891-2010)	0.94	0.92	0.94	0.89
cr(1891-1950 , 1951-2010)	0.48	0.75	0.78	0.70

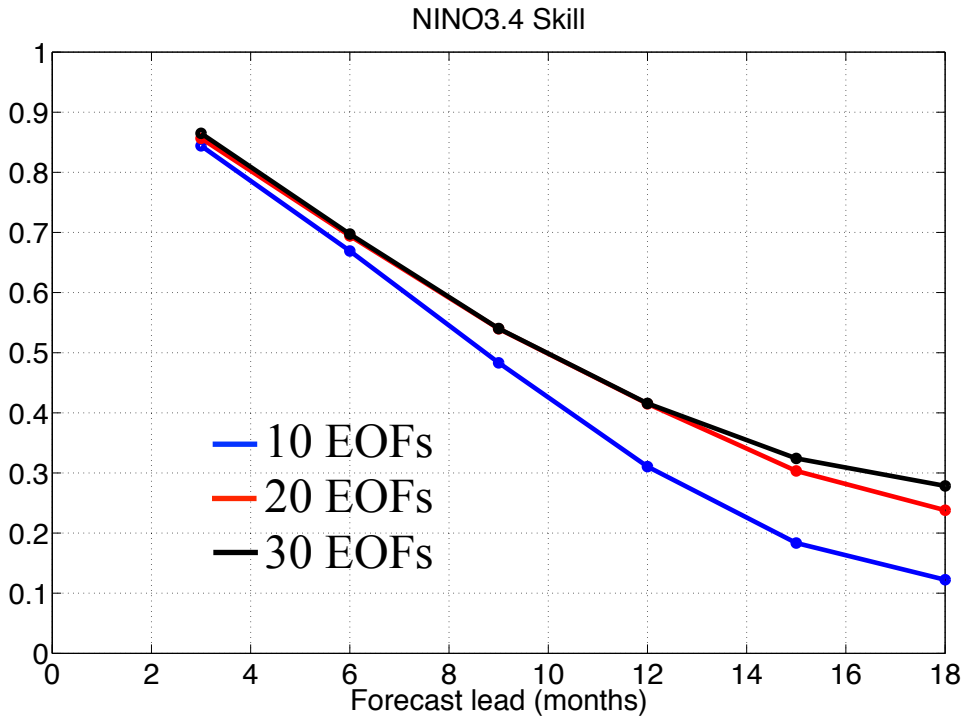
Table 1: Pattern correlations between dominant optimal initial conditions for LIMs constructed for the three periods 1891-1950, 1951-2010, and 1891-2010 using $\tau_e = 3$ months.



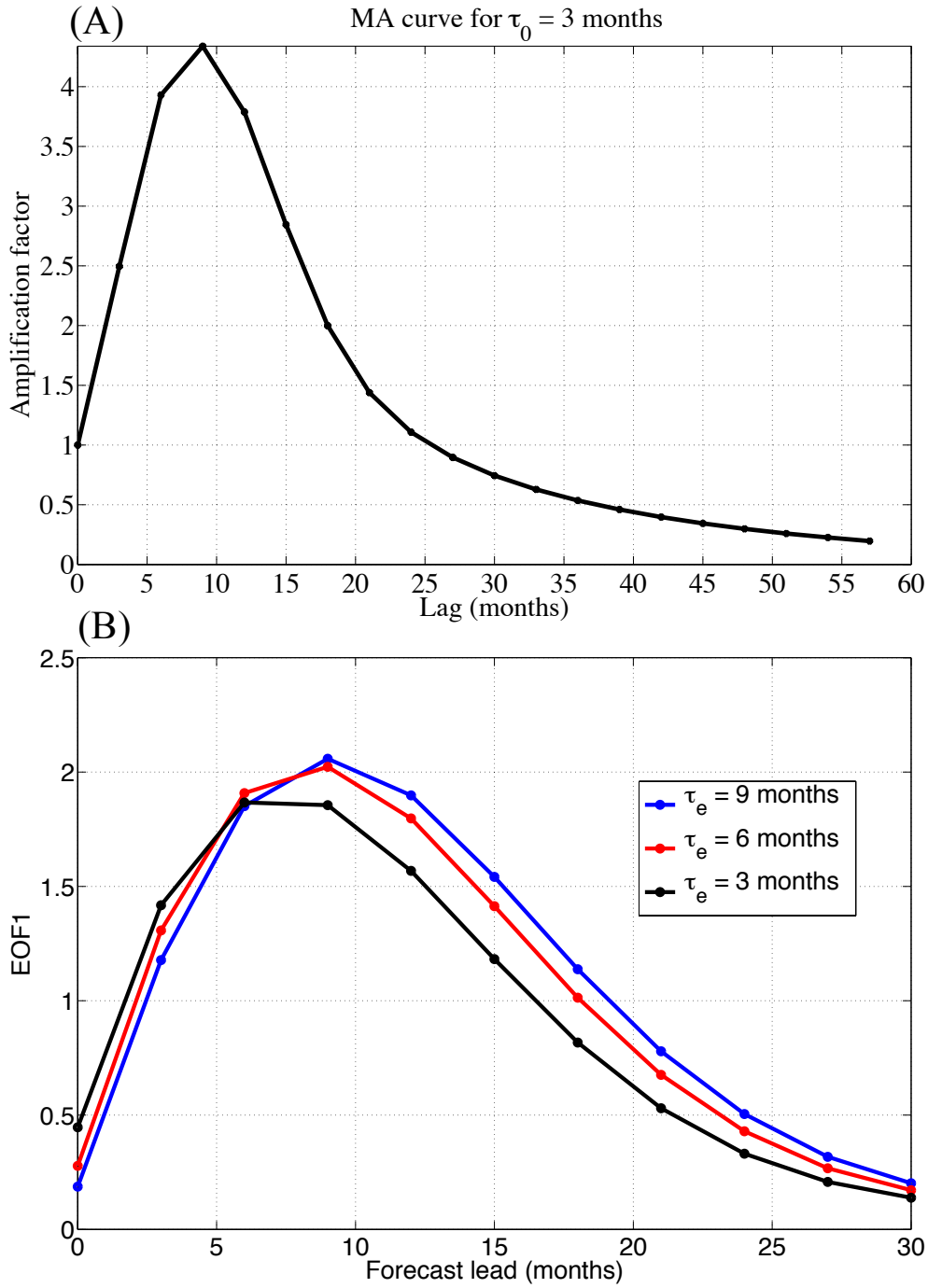
Supplemental Figure 1: SODA nondimensional dominant optimal initial SST anomalies, regressed to ocean temperature structure and surface wind stress anomalies, for τ_e equal to 3 months. Contour interval equal to 0.2. A) At 5m. B) Along the equator. Thick blue contour marks the depth of the climate mean 20°C isotherm. C) At 160°W.



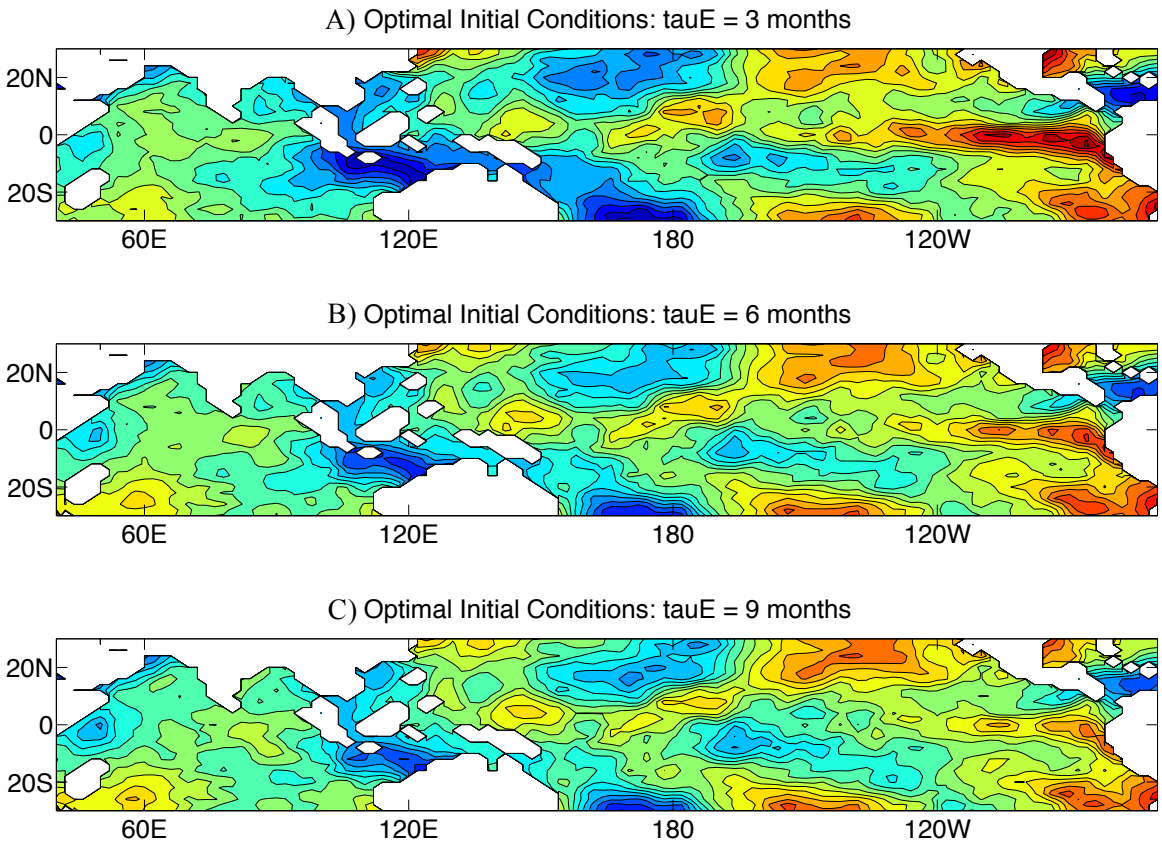
Supplemental Figure 2: Lag-covariance using HadISST seasonal mean 1891-2010, in units of °C^2. A) Observed 9-month. B) LIM 9-month. C) Observed 18-month. D) LIM 18-month. Contour interval is 0.02 °C^2.



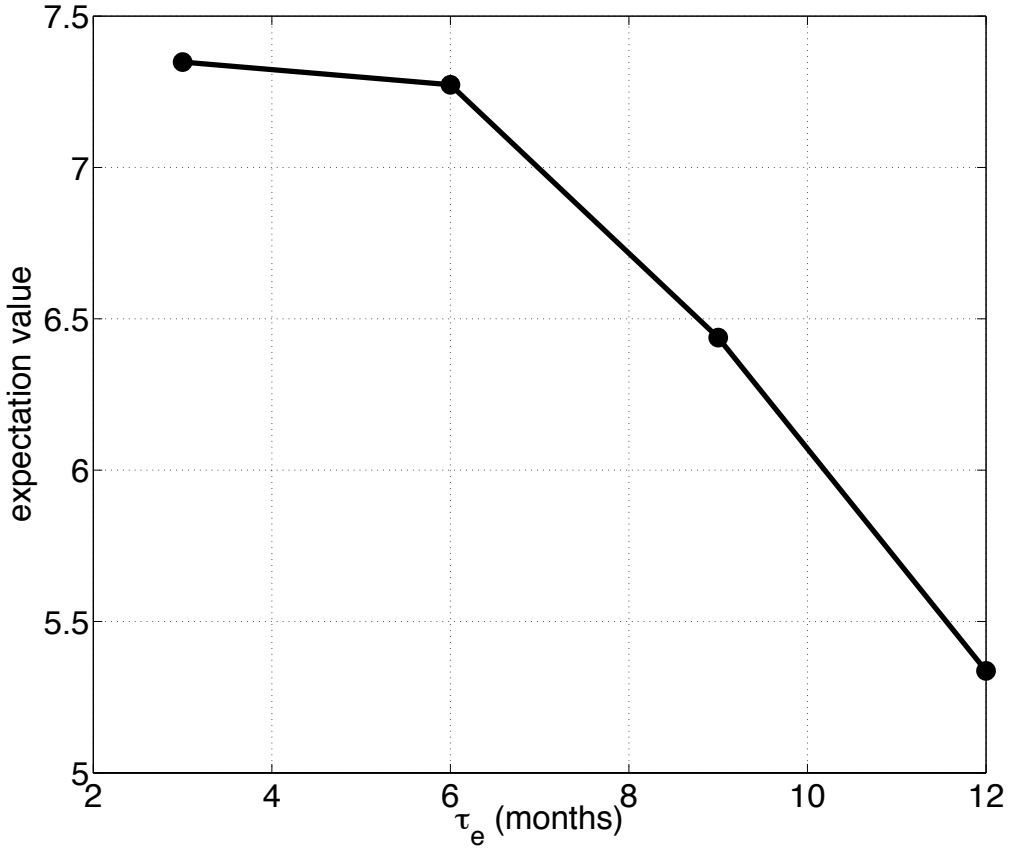
Supplemental Figure 3: Cross-validated forecast skill of HadISST (1891-2010) seasonal mean SSTs averaged over the NINO3.4 region (5°S–5°N, 170°W–120°W) as a function of EOF truncation for LIM with $\tau_o = 3$ months.



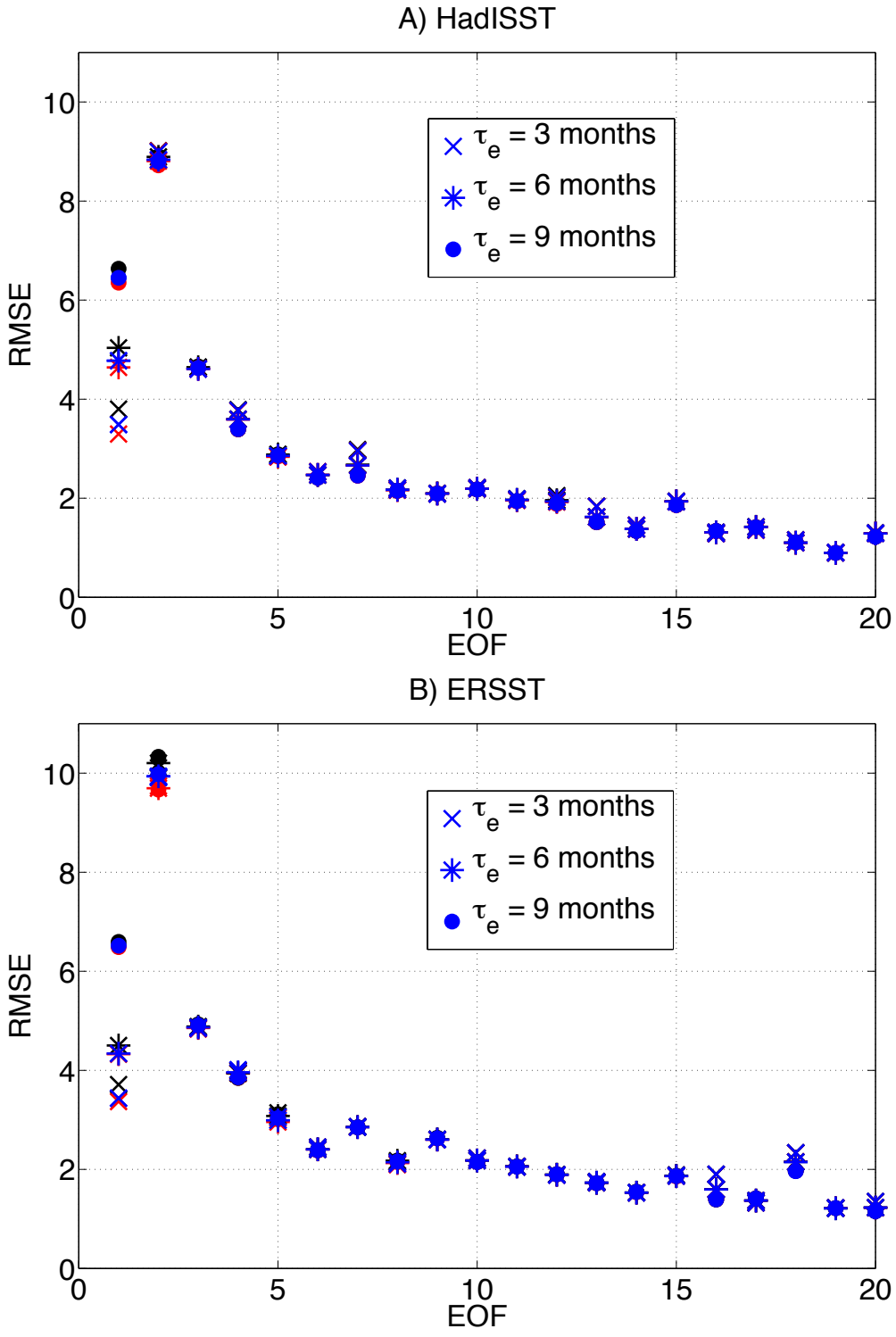
Supplemental Figure 4: A) The largest SST amplification possible as a function of lag--the maximum amplification (MA) curve, determined by a singular value decomposition of G under the L2 norm of SST. B) Projection of first right singular vector on EOF1 for τ_e=3,6,9 months and forecast for 3-30 months, in units of °C. Calculated using seasonal mean HadISST SSTs over the 1891-2010 period.



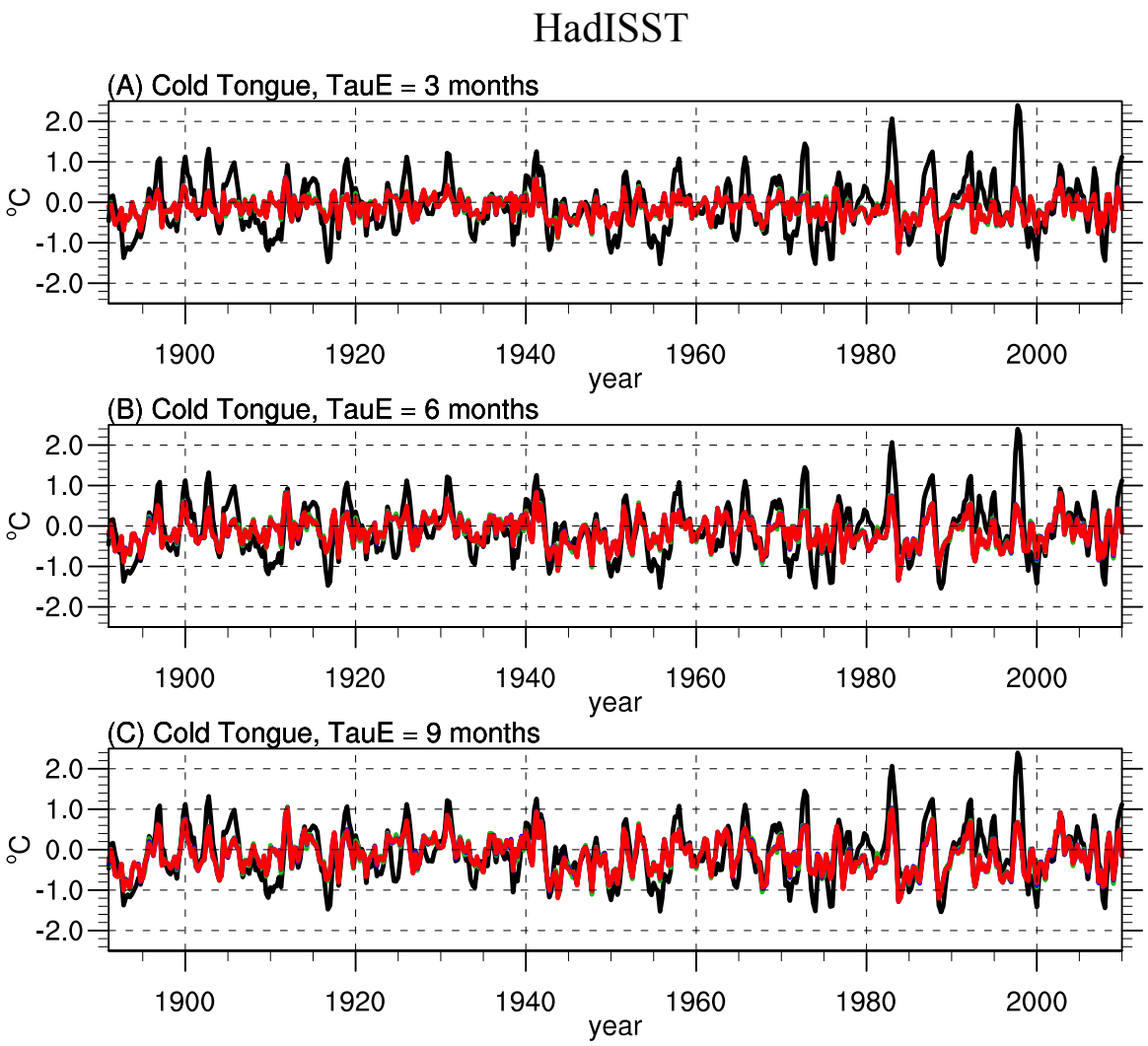
Supplementary Figure 5: Optimal initial conditions (nondimensional), constructed with statistics from 1891-2010 HadISST SSTs. (A) $\boldsymbol{\tau_e = 3}$ months. (B) $\boldsymbol{\tau_e = 6}$ months. (C) $\boldsymbol{\tau_e = 9}$ months.



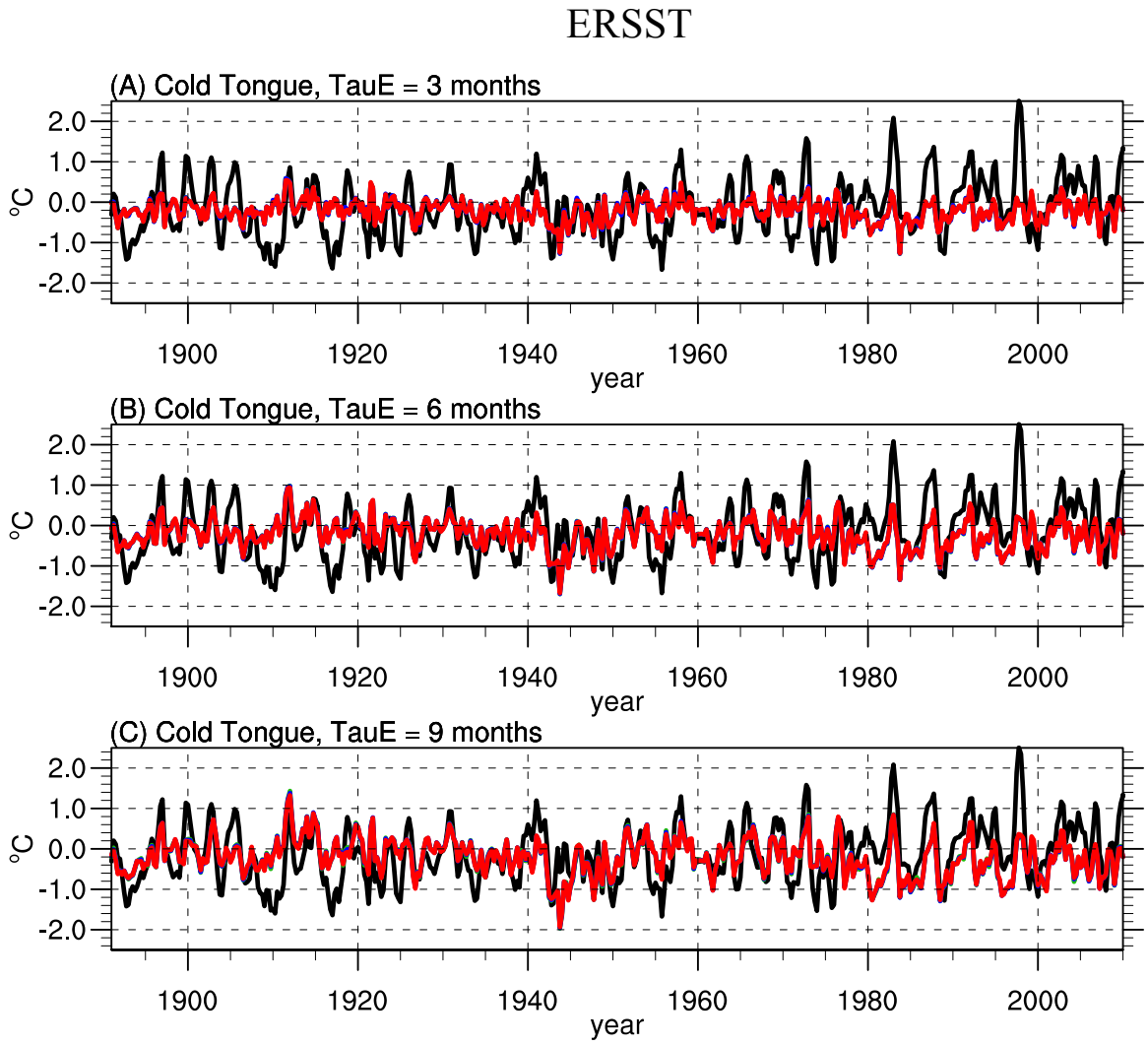
Supplementary Figure 6: Expectation values as a function of $\boldsymbol{\tau_e}$, using HadISST 1891-2010 seasonal mean SSTs.



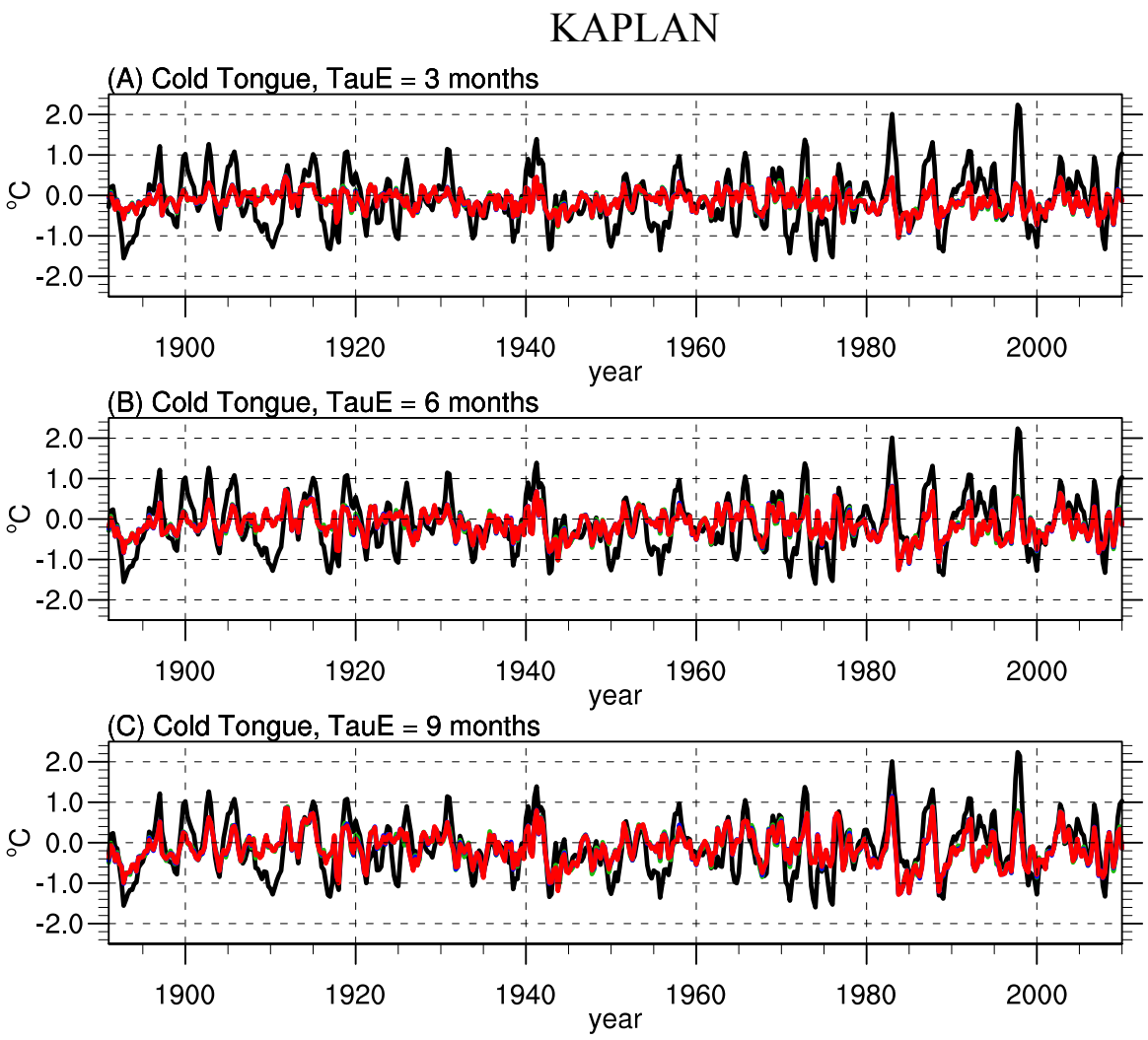
Supplemental Figure 7: A) RMS of filtered data minus forecast error, $\mathbf{RMS}(\mathbf{R}(\boldsymbol{\tau_e}) - \boldsymbol{\sigma}(\mathbf{t}))$ where the filter removes the evolution of the first right singular vector for τ_e = 3,6,9 months and forecast for 21,24,27 months (shown with black, blue, red, respectively) as a function of EOF, in units of °C. For reference, the unfiltered RMSE for EOF1 = 14.3°C and EOF2 = 9.4°C. Calculated using seasonal mean HadISST SSTs over the 1900-2010 period. B) Same as A) for ERSST.



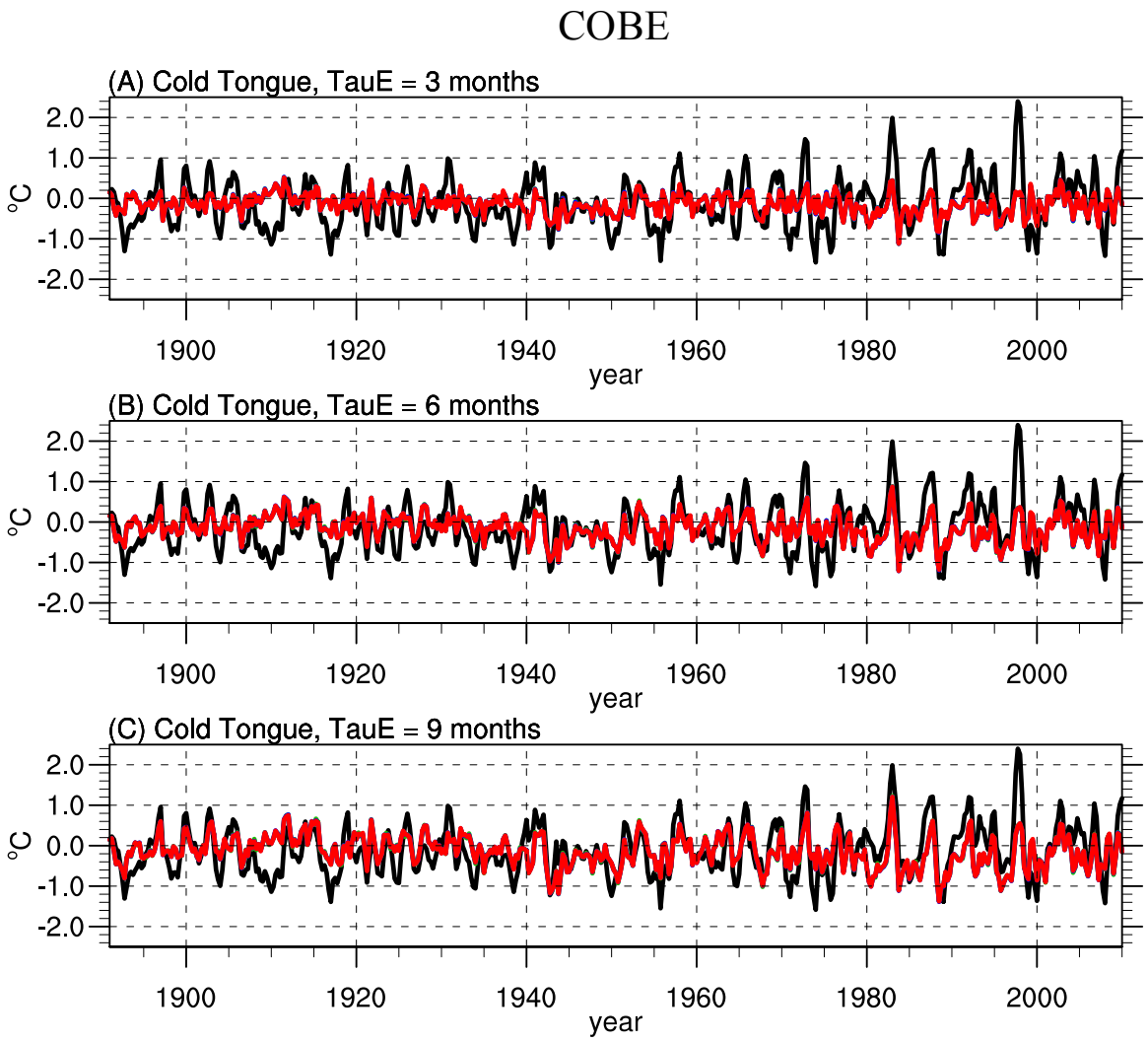
Supplemental Figure 8: HadISST seasonal mean 1891-2010 cold tongue SSTs (defined as the average over 4°N-4°S, 170°E-70°W) unfiltered (black), filtered $\boldsymbol{\tau_1 = 21}$ months (green), filtered $\boldsymbol{\tau_1 = 24}$ months (blue), filtered $\boldsymbol{\tau_1 = 27}$ months (red), in units of °C. (A) $\boldsymbol{\tau_e = 3}$ months. (B) $\boldsymbol{\tau_e = 6}$ months. (C) $\boldsymbol{\tau_e = 9}$ months.



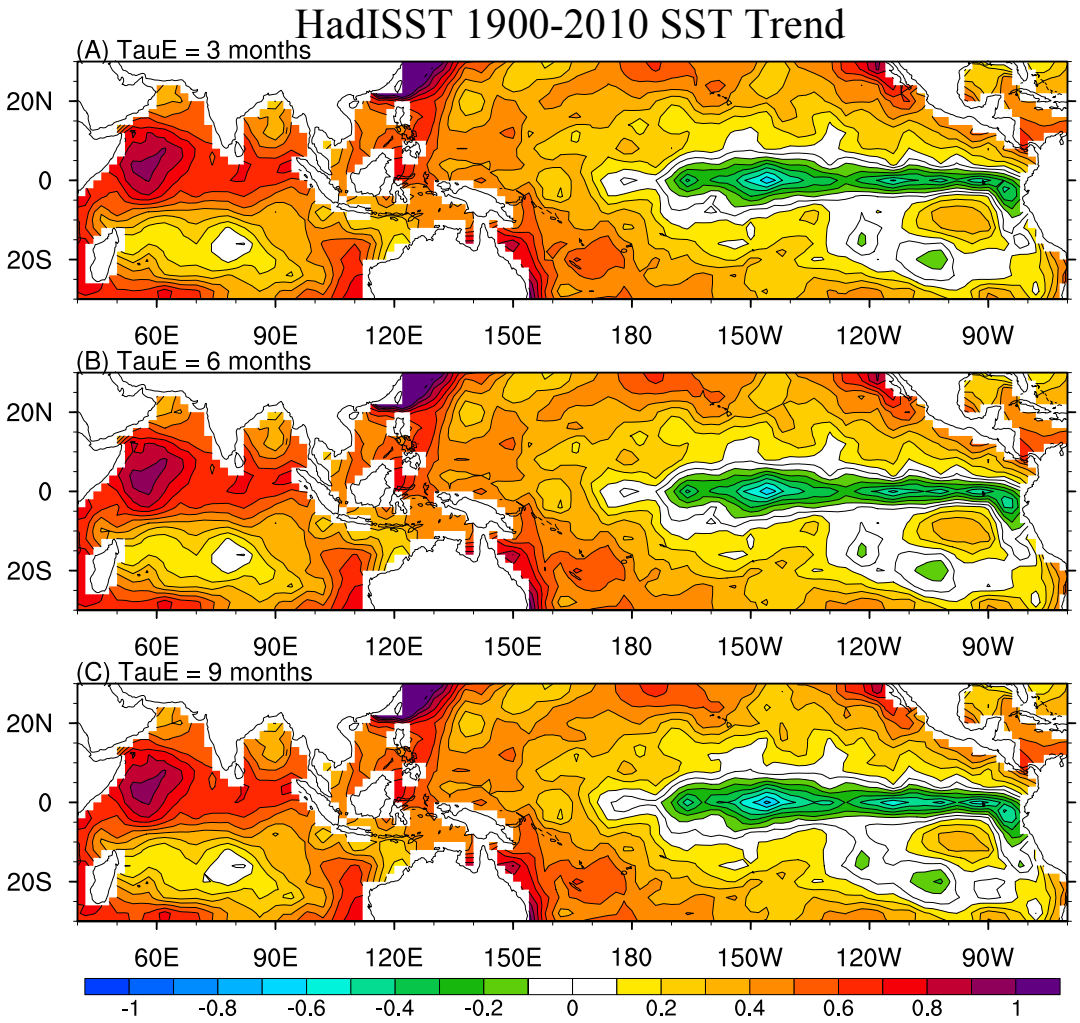
Supplemental Figure 9: Same as Supp. Fig. 8 for ERSST.



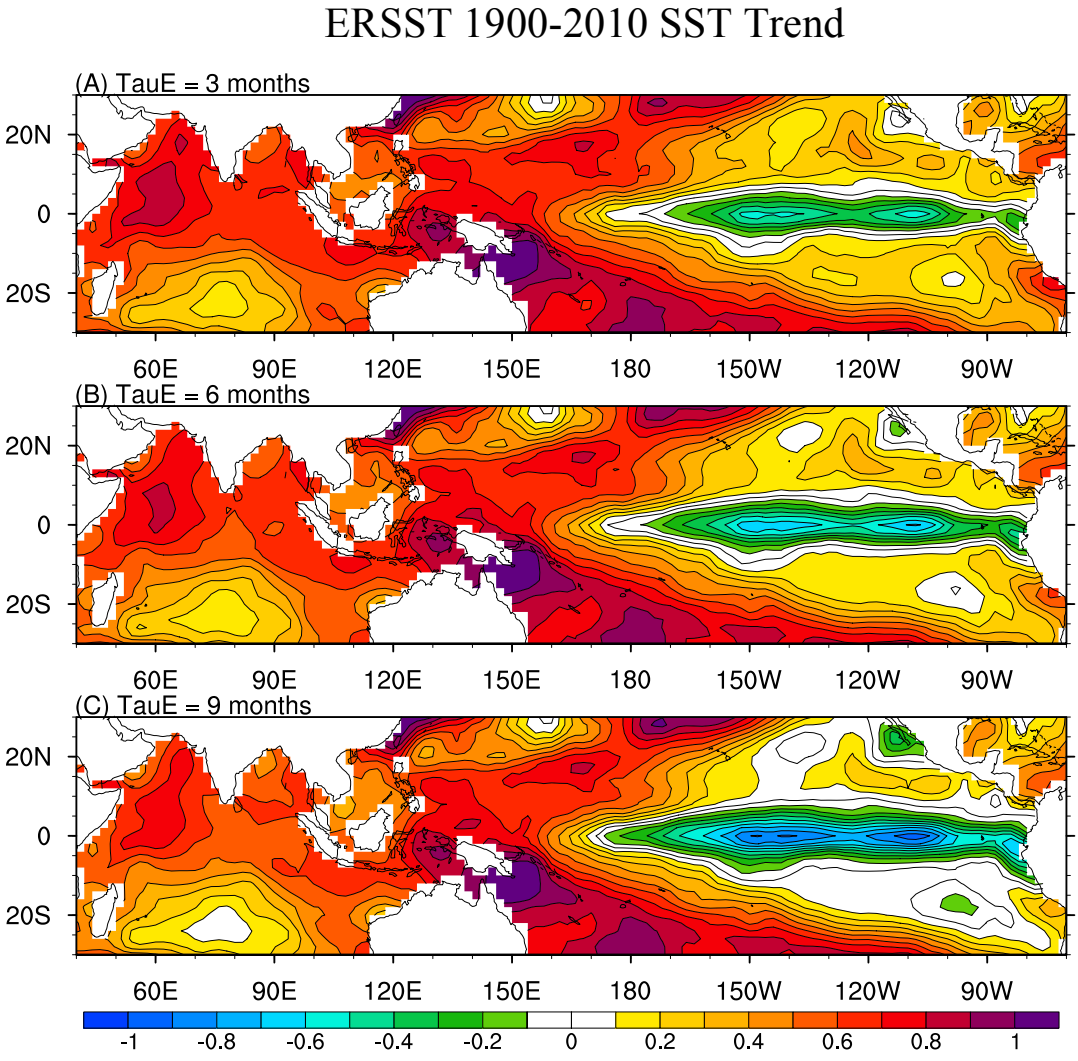
Supplemental Figure 10: Same as Supp. Fig. 8 for KAPLAN.



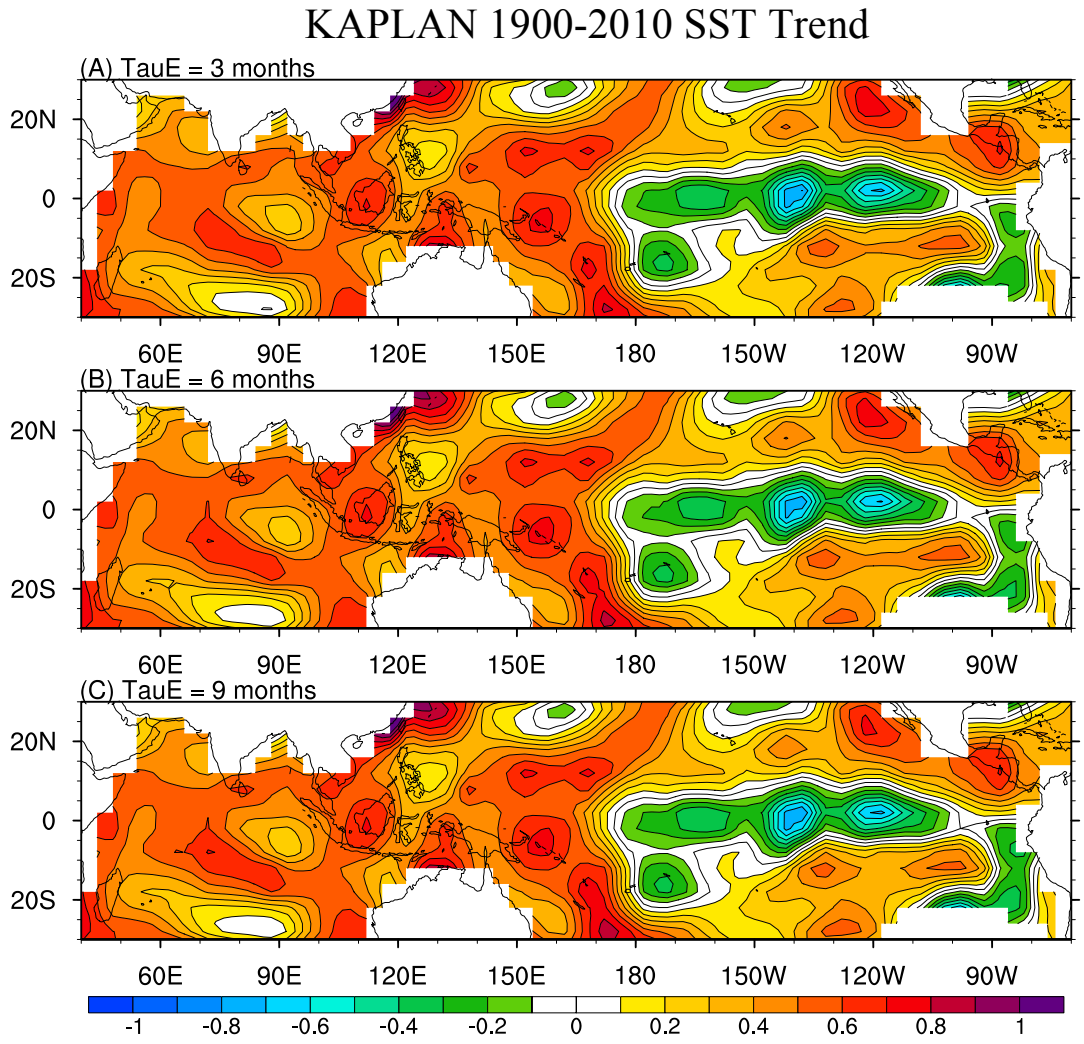
Supplemental Figure 11: Same as Supp. Fig. 8 for COBE.



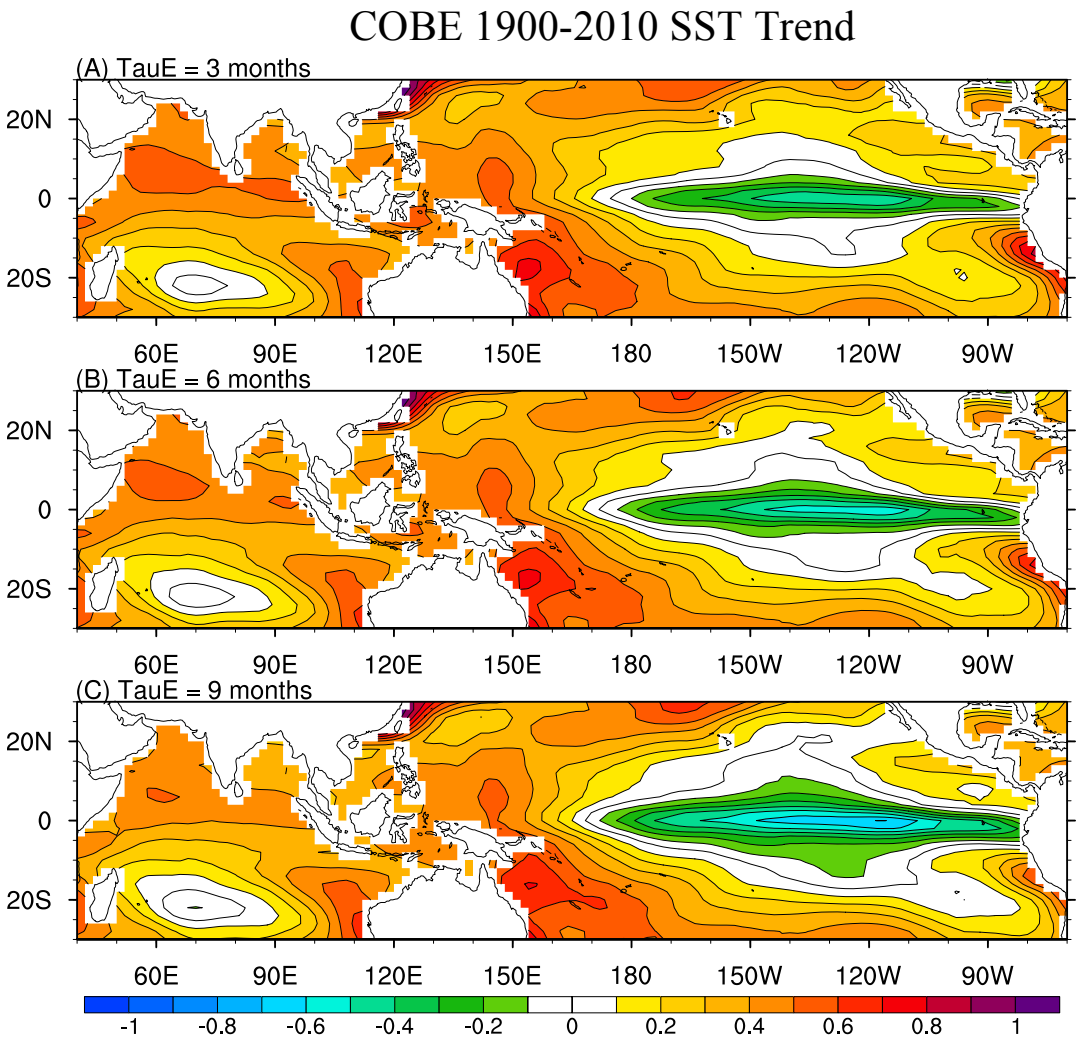
Supplemental Figure 12: HadISST 1900-2010 cold tongue SST trend for $\boldsymbol{\tau_1 = 21}$ months, in units of °C/100 years. (A) $\boldsymbol{\tau_e = 3}$ months. (B) $\boldsymbol{\tau_e = 6}$ months. (C) $\boldsymbol{\tau_e = 9}$ months.



Supplemental Figure 13: Same as Supp. Fig. 12 for ERSST SSTs.



Supplemental Figure 14: Same as Supp. Fig. 12 for KAPLAN SSTs.



Supplemental Figure 15: Same as Supp. Fig. 12 for COBE SSTs.